A New STAR Event Reconstruction Chain

C. A. Pruneau, A. Rose, Wayne State University
M. Calderon, J. Lauret, Brookhaven National Laboratory,
B. Hippolyte, Yale University,
STAR Collaboration.

Abstract

We present the design and performance analysis of a new event reconstruction chain deployed for analysis of STAR data acquired during the 2004 run and beyond. The creation of this new chain involved the elimination of obsolete FORTRAN components, and the development of equivalent or superior modules written in C++. It features new modules for fast TPC cluster finder, track reconstruction (ITTF discussed at CHEP2003) integrating all detector components of the experiment, vertex finding, and various post-tracking analysis tasks including V0 and track kink finding.

Introduction

STAR is a multipurpose detector designed for the study of heavy-ion and p-p collisions at the Relativistic Heavy Ion Collider (RHIC) [1]. Based on a large half Tesla field solenoidal magnet, it comprises a number subsystems including a central time projection chamber (TPC), a vertex detector (SVT, SSD), two forward time projection chambers (FTPC), barrel and end-cap calorimeter systems. The TPC forms the main body of the detector and provides full azimuthal coverage in the rapidity range $|\eta|<1.5$. The FTPCs extend this coverage to $|\eta|<3.5$. The vertex detector consists of three layers of silicon drift chambers (SVT), and one layer of silicon strip sensors (SSD). Observed particle multiplicities range from a few 10s, in p-p collisions, up to 6000 in central Au + Au collisions. Zero suppressed data readout amounts to approximately 20 MB for Au + Au central collisions. The data acquisition system was upgraded to enable a data throughput of 150 MB/s. In its last run, STAR accumulated respectively 15×10^6 and 94×10^6 Au+Au events at sqrt(s_{NN}) of 62 GeV and 200 GeV, and over 200 millions p + p collisions at sqrt(s) = 200 GeV. These amount to over 230 TB of raw data.

The analysis of these data proceeds in two broad stages. The first stage consists of the reconstruction of raw digitized detector data into detector hits, energies, particle trajectories, momenta, and species identity, cumulating into the production of data summary files consisting of list of reconstructed particles, and their properties.

The second stage involves detailed analyses of the reconstructed collisions for study of specific physics phenomena and observables.

The first stage is based on a reconstruction program called "big full chain" (BFC). The BFC essentially consists of a long sequence of calls to "maker" modules performing specific data reduction tasks such as a calibration, hit finding, and track reconstruction. The BFC is a collage of components written in FORTRAN, C, and C++. Track finding and Kalman fitting maker modules were initially implemented as separate maker components (TPT and EGR) to handle track reconstruction of TPC hits. As new detectors were integrated, additional maker modules needed to be integrated to handle track extension through these modules. We observed that the reconstruction time increased linearly with the number of maker modules added to the chain rather than the number of added hits. Already in 2000, it became increasingly clear that having a mix of components written in FORTRAN, C, and C++ rendered the maintenance of the BFC tedious and difficult. Key modules were poorly documented, and some of the developers had or were about to leave the collaboration. The Collaboration thus decided to undertake a re-write of the track reconstruction modules and selected components of the BFC.

New Track Finding/Fitting Module

An integrated tracker task force (ITTF) was appointed in Nov 2000, with the task of designing, developing, and deploying a new tracker, which enabled seamless integration of existing and future detector components. The new tracker was required to enable track reconstruction matching or surpassing that of the older tracker. To be written exclusively in C++, the new software was also required to be platform independent, with minimal dependencies on 3rd party libraries and software components, to be easily maintainable and extendable, and be fully documented. The new tracker design was approved through an internal review in Sept 2002. The full implementation and tuning were completed during the following year. Results were present at CHEP 03 [2]. The tracker was then subject to a performance review in Aug 2003.

The performance review was based mainly on comparative analyses of the old and new trackers for reconstruction of simulated data consisting of HIJING event, processed with GEANT, and detailed STAR detector response simulators. The review focused mainly on the tracker performance based on TPC hits only given this detector's performance was best understood at the time of the review. The two trackers were found to have equivalent track reconstruction efficiency, and identical momentum resolution characteristics. Reviewers thereby recommended the new tracker be adopted for future data reconstruction. This recommendation was officially adopted by the collaboration at its Aug 2003 meeting.

New Reconstruction Chain

The integration and deployment of the new tracker in the STAR BFC required various BFC modules and components be modified or upgraded. An integration team was thus formed to conduct these tasks. We herein summarize the modifications which were done in order to deploy the new tracker, and which to a great extent proceeded in parallel.

The BFC was originally written with tools inherited from the STAF environment, and included table structures accessible in FORTRAN, C, and C++. It emerged that the maintenance of these codes was rather fastidious and time consuming. The decision was then taken to eliminate tables from the BFC, and to use C++ constructs only to store data. It was also required that the STAR persistent event model, "StEvent", become the reconstruction chain backbone for data storage. Multiple maker modules thus had then to be eliminated, modified, or moved within the BFC. New options were also added to the chain in order to enable more dynamic management of the chain thereby enabling event reconstruction based on dynamic STAR detector subsets. New maker modules were created for primary vertex finding, TPC hit cluster finding, TPC hit position calibration, and Kink finding. Various other modifications were also brought to maker modules associated with the SVT, EMC, and FTPC detectors. Performance assessment modules were modified and created to enable a detailed comparative analysis of the old and new chains.

Once the new maker modules were developed and deployed within the new BFC, it became imperative to conduct a full performance analysis of the new reconstruction chain. A team of testers composed of representatives of all STAR physics working groups was assembled for this purpose. Their mandate was to assess the integrity of the data produced by this new chain, and conduct a comparative performance analysis of the old and new chains for specific physical observables of interest. This analysis was conducted on large datasets of d + Au and p + p collisions acquired in 2003. We provide, in the following, key examples of this comparative analysis.

Given the primary vertex finder modules were essentially re-written from scratch in C++, much attention was dedicated to a comparative study of the vertex finding efficiencies achieved with the two trackers. As a specific example, Figure 1 displays a plot of the event vertex yield obtained with the old (black) and new (red) chain as a function of the vertex position along the beam direction. While equal number of raw events was processed with the two chains, one finds the yield obtained with the new chain is slightly higher and thus conclude the new chain has a slightly better vertex finding efficiency.
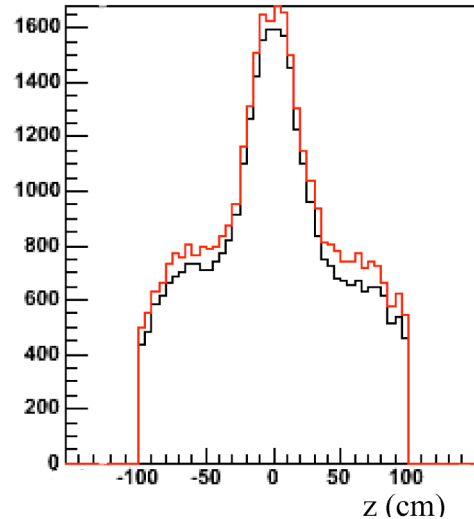

z (cm)

Figure 1 Comparison of the vertex position distributions obtained with the old (black) and new (red) chains.

The strangeness physics working group conducted a comparative analysis based on the yield of reconstructed Λ baryons. The reconstruction is based on a topological analysis of track pairs forming "v0" candidates. Figure 2 presents invariant mass spectra obtained with the old (red) and new (black) chains while processing equal numbers of events. One finds the two spectra have comparable combinatorial background levels and exhibit Λ peaks with essentially the same position, width, and amplitude. A quantitative analysis reveals the new chain in fact seem to outperform the yield old chain by
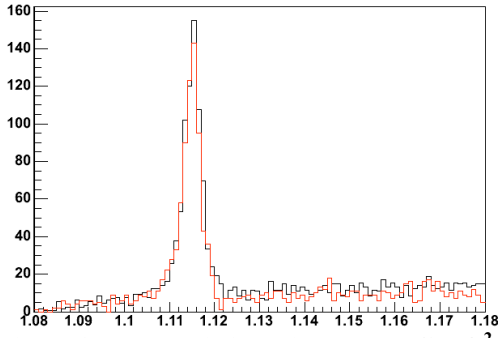
approximately 10%.

Figure 2 Track pair invariant mass (GeV/c^2) spectra obtained with the old (red) and new (black) chains.

The HBT working group conducted a comparative analysis of the track reconstruction quality achieved with the two chains. Figure 3 displays a plot of the normalized number of track pairs as a function of their average separation in the TPC.
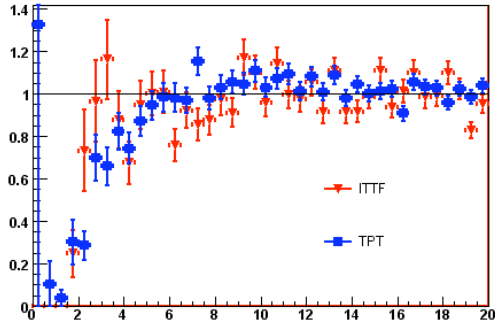

Figure 3 Average two-track separation (in cm) observed in the TPC with the old (blue) and new (red) chains.

Yields fall at low average separation in part because of merging occurring in both hit and track reconstructions. The plot illustrates that within the statistics afforded with this data sample, the two trackers exhibit similar merging levels in the hit and track reconstruction.

Figure 4 displays a comparison of the level of track splitting produced by the two chains. Splitting arises when the track reconstruction software shares the hits of a real track between two reconstructed tracks. The pair yield is plotted as a function of the split quality and the invariant momentum difference of track pairs. The quality ranges from -0.5 (no splitting) to 1 (maximum splitting). Observe that the old tracker produces many split tracks at low invariant momentum difference whereas the new tracker produces comparatively very few.
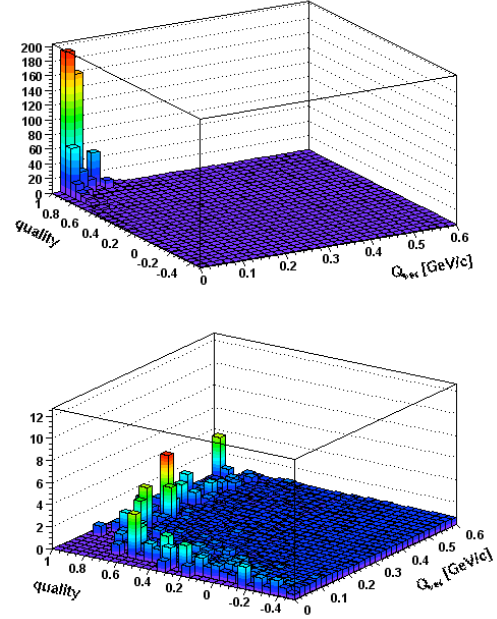

Figure 4 Track splitting distributions obtained with old (top) and new (bottom) chains.

Many other performance metrics were used to insure the production quality of the new reconstruction chain. The event-by-event group found observables used in the analysis of directed and elliptical flow including azimuthal reaction plane angle distributions, raw "v1" and "v2" coefficients, have identical values within statistical errors. The event-structure group likewise found the 2-particle correlations in rapidity and momentum space were equivalent. The spectra group considered vertex distributions, total multiplicity distributions, and particle spectra and found excellent agreement between the two chains. The spectra group also conducted detailed analyses of the particle identification variables based on differential energy loss (dE/dx) measured in the TPC and established the two chains have equivalent particle identification performance as well. The photon yields measured on the basis of conversions within the detector (e.g. in the beam pipe, the SVT, or the inner field cage of the TPC) were found to agree within 5%. Performance comparisons were additionally conducted on the basis of p + p and Au + Au data acquired at sqrt(s_{NN})=200 GeV.

Overall, the working group reported that they were satisfied with the performance of the new reconstruction chain and tracker. The new chain was therefore approved for production. Production of the Run 4 (2004) data reconstruction has now begun. Given the extremely large volume of data acquired during Run 4, it is estimated the production

will require in excess of 12 months at the BNL/RCF reconstruction farm. We remark in closing that STAR has developed, over the years, an elaborated insurance quality system to control and monitor the quality of the data reconstruction. Needless to say this system will be used to systematically insure the quality of this new production.

Concluding remarks

The development, deployment, and performance evaluation of the new tracker and reconstruction chain is completed and production has now begun with the new chain. A large effort was expanded to validate the performance of the new chain and insure the robustness of the code as well as the physics results it produces. The new chain is found to be indeed robust, memory leak free, with solid exception handling, and with equal or superior physics performance. It is also appreciably faster than the previous chain, and given its modular plug-and-play design for track reconstruction, detector configuration and processing, the new tracker and reconstruction chain emerge as a powerful and flexible tool with excellent growth potential to serve the data reconstruction needs of the STAR collaboration well into the next decade.

References:
1. Nuclear Instrumentation and Methods A 499, 624 (2003) and references therein.
2. M. Calderon, J. Lauret, M. Miller, B. Norman, C. Pruneau, A. Rose, Proceedings of the Computing in High Energy and Nuclear Physics (CHEP03) Conference, La Jolla, California, March 24-28, 2003.